

La evaluación continua: una aportación a como corregir algunos efectos negativos

The continuous evaluation: a contribution to how to correct some negative effects

Joan Domingo, Joan Segura, Antoni Grau

Departamento de ESAII (Enginyeria de Sistemes, Automàtica i Informàtica Industrial)
UPC, Universitat Politècnica de Catalunya
Campus Diagonal Sud, Edifici H. Av. Diagonal, 647 08028 Barcelona
joan.domingo@upc.edu joan.segura@upc.edu antoni.grau@upc.edu

Resumen

La evaluación continua supone que el estudiante debe realizar una importante cantidad de actividades que, en la mayoría de casos, se traduce en calificaciones que le orientan, junto con las observaciones realizadas por el docente, sobre el resultado de dichas actividades, a fin de encaminar al estudiante hacia el aprendizaje pretendido. En la evaluación continuada, una de las acciones posibles es permitir al estudiante recuperar las actividades no superadas, enmendando sus errores, a fin de que la calificación pase a aprobado. Así, cada estudiante acaba obteniendo a lo largo del curso una gran cantidad de calificaciones, con lo que la media obtenida será mayoritariamente de aprobado. La ley de los grandes números y el teorema del límite central apuntan, cada uno a su manera, que cuando la cantidad de calificaciones de cada estudiante es suficientemente elevada, su distribución tiende a una distribución normal. Por ello, la evaluación continuada ha sufrido algunas observaciones negativas en relación a que no discrimina suficientemente a los estudiantes a partir de sus calificaciones finales y, a la vez, no permite que los estudiantes más brillantes obtengan calificaciones excelentes. En este artículo, se prueba que, efectivamente, la distribución de calificaciones en un sistema de evaluación continuada sigue aproximadamente una distribución normal, y se establece un mecanismo de corrección basado en un polinomio de calificación diversificado y de una función correctora.

Palabras clave: Evaluación continuada, Función correctora, Ley de los grandes números, Teorema del límite central.

Abstract

The continuous evaluation supposes that the student must realize an important amount of activities that, in most of cases, is translated in grades that orient students, along with the observations realized by the teacher, on the result of said activities, in order to direct the student towards the intended learning. In the continuous evaluation, one of the possible actions is to allow the student to recover the unsatisfactory activities, amended their errors, so that the qualification passes to approved. Thus, each student ends up obtaining a large number of grades throughout the course, which means that the average obtained will be mostly approved. The law of large numbers and the central limit theorem, each in its own way, points that when the number of grades of each student is sufficiently high, its distribution tends to a normal distribution. Therefore, the continuous evaluation has suffered some negative observations in relation to his inability to sufficiently discriminating students from their final grades and, at the same time, does not allow the brightest students to obtain excellent grades. In this paper, it is proved that, in effect, the distribution of ratings in a continuous evaluation system follows approximately a normal distribution; also, a correction mechanism based on a diversified qualification polynomial and a corrective function is established.

Key words: Continuous evaluation, Corrective function, Law of large numbers, Central limit theorem.

1. Introducció

La evaluació en el àmbit acadèmic responde, al menys, a tres tipus bàsics sin detriment de qualquier otra calificació:

- *Evaluación diagnòstica*: para determinar el nivel de conocimientos previos de los estudiantes sobre un tema.
- *Evaluación formativa*: también denominada evaluación auténtica, por cuanto mide los conocimientos de los estudiantes y, a cada uno de ellos, se le hacen las observaciones oportunas para corregir deficiencias; exige una atención a la diversidad puesto que no todos los estudiantes cometen los mismos errores ni los cometen en la misma cantidad.
- *Evaluación sumativa*: también denominada evaluación calificativa, es la que otorga una valoración, generalmente numérica, sobre el grado de consecución porcentual de los conocimientos que se espera que un estudiante haya alcanzado en un período de tiempo académico o sobre un determinado tema o conjunto de temas. No informa al estudiante de sus errores, sólo cuantifica su nivel de acierto.

Tradicionalmente se ha considerado, como señalan Jorba y Sanmartí (Jorba, J.; Sanmartí, N., 1994) que la evaluación diagnòstica se realiza antes del proceso de enseñanza-aprendizaje, que la formativa se realiza durante el mismo y que la sumativa se realiza después. Nuestro enfoque consiste en mezclar las dos últimas durante el proceso de enseñanza-aprendizaje a fin de poder dar una mejor información al estudiante, a la vez que mejorar en cada momento su progreso académico. Dentro de esta evaluación continuada formativa-sumativa, se pueden dar dos variantes: la que no permite recuperar calificaciones y la que sí lo permite. Entendemos que no permitir que un estudiante demuestre que ha aprendido de sus errores es un camino equivocado, puesto que el estudiante accede a los estudios sin conocimientos sobre las materias que allí se imparten y, por tanto, es normal que cometa errores. Así, entender que el error en el grado de conocimiento de las materias es muy elevado o incluso completo, supone aceptar que, a media que se vayan realizando actividades, este error debe ir disminuyendo o lo que es igual, que el aprendizaje debe ir aumentando.

Tenemos por cierto que, según establece Hervás (Hervás, 2008), no todos los estudiantes son iguales ni aprenden al mismo ritmo ni requieren del docente las mismas ayudas. Siguiendo a Alonso, Gallego y Honey (Alonso, Gallego y Honey, 1994) los alumnos deben “*aprender a aprender*” y “*los profesores deben reconocer las diferencias individuales de sus alumnos para personalizar su educación tratando de que sus preferencias en cuanto a los Estilos de Enseñanza no influyan en los Estilos de Aprendizaje de los alumnos*”. García Cué (García Cué, 2006), que coincide con Alonso, Gallego y Honey, añade que “*los profesores enseñan como les gustaría aprender las cosas*”. Además, como ya establecieron Willis y Hodson en 1999, “*para la mayor parte de los alumnos no hay dificultades de aprendizaje en cualquier contenido temático si se les enseña según sus propios estilos de aprendizaje. El éxito en el aprendizaje se basa en la capacidad para ajustar a cada uno de los alumnos su propia forma de aprender.*”

No es que sigamos a Gardner (Gardner, 1983) y a su teoría de las inteligencias múltiples puesto que la comunidad científica no avaló en su momento sus ideas (y sigue sin hacerlo), las cuales tampoco venían avaladas por ningún tipo de evidencia empírica, sino que se basaban en una intuición y no fue hasta 2006, 23 años después de publicar sus ideas, que (Visser y otros, 2006a y 2006b) realizaron el primer estudio empírico sobre inteligencias múltiples, dando resultados negativos y estableciendo que sí se puede hablar de una inteligencia general subyacente a las diferentes habilidades, talentos y capacidades de la persona, desmontando una buena idea que no pasó de serlo pero que no tiene demostración, aunque siga teniendo seguidores. Por ello, y a tenor de los estudios de Visser y otros, no entendemos que sean diferentes inteligencias las que se establecen en los diferentes términos que forman el polinomio de calificación de una asignatura (porcentaje debido a cada parte de la asignatura – teoría, prácticas, etc. –) sino que deben ser diferentes tipos de habilidad los que dichos polinomios midan. Ello no significa que desmerezcamos

las aportaciones de Gardner, premio Príncipe de Asturias de Ciencias Sociales 2011, a la evolución de los modelos educativos, solamente que no nos son de aplicación.

Entendemos que el estudiante sólo puede aprender de sus errores y que estos errores debe señalarlos el docente haciendo las observaciones oportunas sobre el trabajo de cada uno de ellos y expresando la gravedad de dichos errores en forma de calificación, puesto que el estudiante no tiene porqué conocer la gravedad de los errores que comete; también esta calificación supone un mensaje sobre ello, teniendo por objetivo formar e informar al estudiante, por una parte sobre los errores en si mismos y, por otra, sobre la gravedad de los mismos. En la Figura 1, se puede ver esta idea.

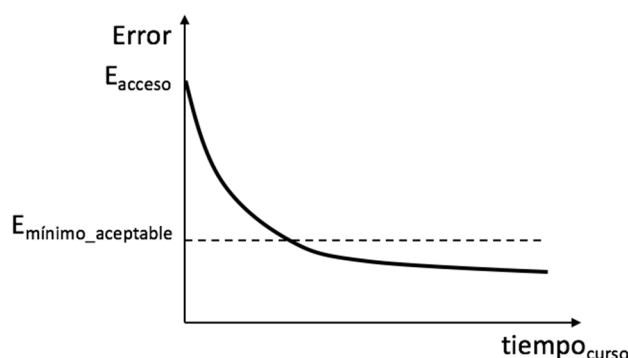


Figura 1: El error en el aprendizaje

El error de acceso (E_{acceso}) es diferente para cada estudiante y es la cantidad de desconocimiento inicial sobre un tema o sobre el conjunto de la asignatura; por ello es importante una atención individualizada que sólo es posible mediante el encargo frecuente de tareas y actividades individuales. El error mínimo aceptable ($E_{\text{mínimo_aceptable}}$) es el que debe alcanzar cualquier estudiante para poder aprobar el curso, y para ello, se dispone de todo el tiempo que dure el curso. Sin duda, habrá estudiantes que superarán con creces este nivel mínimo de error y lo minimizarán más, puesto que el nivel mínimo aceptable, supone una calificación de aprobado con nota del 50% del máximo obtenible y que podrán demostrar valores muy superiores, pero difícilmente podrán demostrar haber alcanzado el 100% en un sistema de evaluación continuada. Cabe señalar que esta misma gráfica, puede extenderse a, por ejemplo, un único tema; en tal caso el tiempo del curso debe considerarse el tiempo de solamente un tema.

Una evaluación continuada que no permita al estudiante demostrar que va superando los errores cometidos es un sistema incompleto, al que le falta también la realimentación necesaria que debe establecerse entre cada docente y cada uno de sus estudiantes para poder comprobar su grado de avance y el grado de satisfacción del mismo. Así, sin este elemento de realimentación permanente que ofrece al estudiante la posibilidad de demostrar que va disminuyendo su grado de error es muy difícil que el estudiante conozca su grado de alcance en conocimientos, lo que equivale a conocer cómo va minimizando sus errores.

La idea, sencilla, sobre esta capacidad del sistema para readmitir las evidencias de los resultados de aprendizaje se puede ver expresada en forma gráfica en la Figura 2.

Las actividades y tareas que se proponen al estudiante se transforman, al ser entregadas al docente, en una suerte de constructos y productos que reflejan el grado de aprendizaje del estudiante en relación a aspectos de forma, de fondo y hasta normativos; la norma sobre cómo, dónde y cuándo presentar los resultados, la forma y formato en qué realizarlos y los contenidos que debe haber, entre otros posibles. Se trata de entrenar al estudiante en el uso de las competencias específicas de una determinada materia y, las competencias, sólo se pueden adquirir con la práctica (y no viendo practicarlas a los demás) y

demostrándolas, en este caso, mediante estos constructos y productos de aprendizaje orientados a la demostración de a qué ritmo se van adquiriendo.

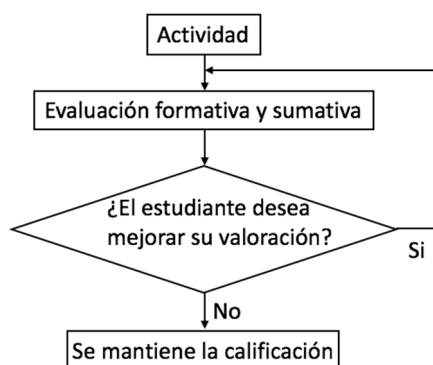


Figura 2: Readmisión de constructos de conocimiento

Un sistema de enseñanza-aprendizaje orientado a las competencias exige, por tanto, un sistema de evaluación continuada y, a la vez, una solicitud muy frecuente de trabajo del estudiante para que pueda poner de manifiesto cómo va adquiriendo las destrezas mínimas requeridas (el error mínimo aceptable). La evaluación continuada es, pues, no sólo poner más tareas al estudiante para tener más calificaciones, sino tener más ocasiones de interactuar con él para, así, mejorar el proceso de aprendizaje tal como proponen (Delgado y Oliver, 2006), *“el profesor no sólo debe evaluar al final del proceso de aprendizaje la asimilación de conocimientos y el desarrollo de competencias por parte de los estudiantes, sino que, a lo largo del curso, debe proponer con cierta periodicidad actividades, de carácter evaluable, que faciliten la asimilación y el desarrollo progresivos de los contenidos de la materia y de las competencias que deben alcanzarse, respectivamente”*.

1.1. Críticas al sistema de competencias

Se encuentran fuertes críticas a la enseñanza basada en competencias en revistas y publicaciones notables como (del Rey y Sánchez-Parga, 2011), (Barnett, 2001), (Boon y van der Klink, 2002), (informes DeSeCo, 2000), (Díaz Barriga, 2006), algunas realmente antiguas, pero que han dejado un importante sedimento en el profesorado, básicamente universitario, de forma que se ha establecido una cierta corriente de rechazo a la evaluación continuada, base sin duda del aprendizaje basado en competencias. El modelo de aprendizaje basado en competencias, ciertamente, se ha convertido en un mono-modelo tal y como ya apuntaban algunos autores (Moreno y Soto, 2005), pero es un modelo que ha venido para quedarse, al menos por mucho tiempo, puesto que está demostrando buenos resultados y posee notables soportes y apoyos. Como todo modelo, tendrá en el futuro su modelo o modelos sucesores, como el de las competencias ha sido sucesor de otros modelos.

Las competencias son algo que se debe aprender y que hay que demostrar haber aprendido haciendo cosas, no explicando cómo hacerlas; por ello, es requisito el saber demostrarlas y para ello se precisa de una muy frecuente consulta a los estudiantes para que puedan poner de manifiesto sus destrezas y, de ahí, la imperiosa necesidad de establecer sistemas de evaluación continuada, entendida como método de trabajo que permita al estudiante, de forma continuada dar cuenta de sus progresos. Este método de trabajo tiene en cuenta tanto el producto de aprendizaje como el proceso seguido para elaborarlo y, además, tiene en cuenta la diversidad de los estudiantes presentes en el aula puesto que asume que no todos tendrán las mismas habilidades ni las mismas capacidades, que no aprenderán igual ni al mismo ritmo, de forma que en la oferta educativa que se hace a los estudiantes, todos puedan poner de relieve en qué destreza destacan y donde tienen sus debilidades y fortalezas.

Un estudiante que no destaque en ningún tipo de oferta de actividad formativa y de aprendizaje, posiblemente suspenderá la materia, pero es importante que al docente no se le pueda pasar por alto ningún estudiante y que no se les ofrezca a todos y a cada uno de ellos alguna vía para poder demostrar su dominio en algún ámbito (saber, saber hacer, saber ser).

1.2. Polinomio de evaluación

Desde que empezó el denominado “Plan Bolonia”, hemos venido impartiendo la asignatura de Control Industrial y Automatización, común a todas las titulaciones de grado en ingeniería de la rama industrial que se ofrecen en la Escola d’Enginyeria de Barcelona Est, EEBE. Por ella han pasado hasta la fecha 3909 estudiantes, contando los actualmente matriculados.

Desde el principio de la misma se ha venido realizando el sistema de evaluación continuada mixta entre formativa y sumativa y ello ha generado una gran cantidad de información que permite ser analizada desde la perspectiva de que la población utilizada para el estudio, consideramos que es suficientemente significativa como para extraer conclusiones generales.

Por otra parte, entendemos que el polinomio de calificación que se utilice para evaluar es fundamental en este tipo de mecanismo de evaluación continuada puesto que debe estar formado por múltiples conceptos. Asumimos que no todos los estudiantes poseen las mismas destrezas, de forma que si todas ellas forman parte del polinomio de calificación en proporción equivalente, mientras unos destacarán con algunas destrezas (cálculo, gráfico, programación, diseño de soluciones, búsqueda de componentes reales, organización de tareas, prácticas, resolución de casos, ejercicios, etc.) otros destacarán en otras, compensando deficiencias con destrezas.

En nuestro caso el polinomio está formado por las siguientes componentes:

- 10% capacidad de búsqueda eficaz de información
- 20% capacidad para resolver un caso de automatización
- 20% capacidad para resolver un caso de control continuo
- 15% capacidad para resolver un proyecto
- 10% capacidad para resolver casos prácticos
- 10% capacidad para explicar elementos que forman parte de la solución de casos prácticos y elementos reales
- 15% para resolver ejercicios (mitad en equipo, mitad individuales)

Ello permite que los estudiantes puedan demostrar sus destrezas en diferentes ámbitos y no correr el riesgo de no superar una asignatura por no haber acertado con la mejor forma de evaluar a cada uno de ellos.

Cuando la calificación final de la asignatura de un estudiante se debe a una cantidad de calificaciones parciales importante, sobre unas 30, entonces es posible que el valor de dicha calificación final tienda a una cierta cantidad previsible. Efectivamente, es prácticamente imposible obtener un cero en todas las calificaciones (excepto si se abandona la asignatura, claro) así como es prácticamente imposible obtener siempre la máxima calificación. Por ello, un sistema basado en la evaluación continua dificulta en extremo la obtención de notas excelentes, dejando a los más de los mejores estudiantes en la zona del notable, mas o menos alto y la mayoría de los demás estudiantes en la zona del aprobado también más o menos alto.

Por lo anterior, se acostumbra a atribuir al sistema de evaluación continuada basada en múltiples evidencias del trabajo del estudiante y sus respectivas valoraciones numéricas, una cierta imposibilidad de discriminar suficientemente a los estudiantes excelentes puesto que la excelencia es difícil de demostrar en todas las ocasiones. Asimismo, también es crítica frecuente que a los estudiantes menos brillantes, este sistema, les ofrece la posibilidad de aumentar sus calificaciones, aunque sólo sea porque pueden hacer muchos trabajos

de calidad mediocre pero suficiente para alcanzar el aprobado, algo que, si la evaluación sólo se basara en exámenes, les pondría en una situación más acorde con sus conocimientos reales.

No obstante, la evaluación continuada permite que todos los estudiantes encuentren ámbitos en los que poner de manifiesto áreas en los que son más expertos, siendo más diestros quizás aplicando procedimientos que resolviendo ejercicios, haciendo la parte práctica con mayor destreza que la teórica, o al revés, demostrando competencias genéricas de valor innegable como el trabajo en equipo, la capacidad de aprendizaje autónomo o el dominio de idiomas, sólo por citar algunas.

1.3. El teorema del límite central

Por su parte, el teorema del límite central establece que la distribución de la suma estandarizada de variables aleatorias independientes que tienen una variancia finita tiende a una distribución normal estándar cuando el número de términos de la suma crece indefinidamente. Esto, que puede parecer muy técnico, se puede resumir en que si existen suficientes datos inconexos (aleatorios) en su distribución, esto es, la probabilidad que se dé uno de ellos, obedece a una distribución gaussiana (con forma de campana de Gauss), que es como se denomina a la distribución normal. O si se quiere, también puede entenderse como que el promedio de las calificaciones obtenidas por cada estudiante se distribuirá más como si fuera una distribución normal a medida que de cada estudiante se tengan más calificaciones. Rehuiremos las matemáticas puesto que el foco de este artículo no es el objeto matemático en sí, sino su utilidad para los fines que pretendemos comprobar y, además, la demostración y base matemática de este teorema es muy conocida y se puede encontrar en múltiples fuentes.

Así, las calificaciones de los estudiantes, que tienen esta característica de aleatoriedad debida a que los estudiantes son muy diversos, ofrece la posibilidad de considerar que, cuando de un estudiante se dispone de suficientes notas, estas, tenderán a una distribución normal. No obstante, al ser aleatorias, no son cantidades completamente independientes puesto que de un mismo tema, un estudiante es cada vez más diestro y por tanto las calificaciones posteriores acostumbra a ser mejores que las anteriores puesto que se ha aprendido de los errores cometidos en ellas, más, cuando se le ofrece la posibilidad de recuperar malas calificaciones. De esta forma, y a partir de este fallo en la independencia de los datos, es difícil establecer que en un sistema de evaluación continuada, las calificaciones tiendan a distribuirse de forma normal.

1.4. La ley de los grandes números

Adicionalmente, la ley de los grandes números, también la podemos simplificar de la siguiente forma: si tiramos una moneda 10 veces al aire puede que las 10 veces salga cara o cruz, pero que si la tiramos al aire 1000 veces, la cantidad de cara y cruz, tenderá a ser la misma y que con infinitas tiradas, saldría tantas veces cara como cruz. Los grandes números no son números grandes sino que se refiere a la cantidad de eventos. Así, podemos considerar, análogamente, que cuantas más notas tenga un estudiante, más tenderán a un cierto valor, en promedio o que existirá una probabilidad de que se obtenga una determinada calificación. Este valor, razonablemente, no será nunca extremo puesto que difícilmente siempre obtendrá calificaciones, como ya se ha apuntado, de cero o de diez y que, a medida que vaya obteniendo calificaciones, irá mejorándolas puesto que al serle comunicadas a la mayor brevedad, podrá establecer mecanismos de corrección que le permitan desplazar hacia la parte alta sus calificaciones más bajas.

En un sistema de evaluación continuada, el estudiante, esencialmente y tal como se ha establecido antes, aprende de sus errores, de forma que un fallo grave en una tarea con calificación le informa, a partir de la propia valoración y de las indicaciones del profesor, de qué debe mejorar y, si mejora la tarea en una nueva entrega, esta calificación negativa se ve incrementada. De esta forma, cualquier estudiante puede aprender de sus errores y de la guía orientadora del profesor y tener calificaciones que, sin ser excelentes, sí pueden alcanzar la zona del aprobado y del notable. Por ello, cada estudiante dispone de muchos números que son

representativos de sus calificaciones y, sin ser grandes números, por lo que a un sistema de evaluación se refiere, se puede considerar que son suficientemente grandes. Las calificaciones académicas no son realmente eventos aleatorios puesto que sufren una posibilidad de modificación (ejecutadas bajo esta práctica de posibilidad de revisión referida en la Figura 2) pero considerando la gran diversidad de estudiantes, sí podríamos aproximar dicha aleatoriedad.

La posibilidad de revisión obedece a que sólo se aprende de los errores en el momento que se es consciente de ellos y de cuál debe ser su corrección; esta posibilidad de recuperación conduce realmente al aprendizaje mientras que una evaluación negativa sin posibilidad de recuperación sólo conduce a la frustración y a la desmotivación, algo absolutamente contrario a la necesaria motivación que cualquier forma de aprendizaje requiere. Asimismo una calificación elevada pero no excelente, puede dejar al estudiante desinformado sobre qué es lo que le faltaría haber hecho o en qué no ha sido suficientemente acertado para ser excelente.

2. Hipótesis

La hipótesis a verificar en este artículo, es que cuando la calificación final obtenida por un estudiante en una asignatura se debe a un reducido número de calificaciones, la cantidad de matriculados que la aprueban es menor que cuando se debe a un elevado número de calificaciones y que si se tiene un elevado número de calificaciones es posible predecir, hasta cierto punto, la probabilidad de que un estudiante obtenga una determinada calificación.

Se considera que, con los valores obtenidos con el teorema del límite central, no se puede establecer, de antemano, cuáles serán los resultados académicos de un curso puesto que la respuesta de este cálculo siempre se asimilará a una respuesta de tipo normal que no discriminará modelos de evaluación, pero que la ley de los grandes números, sí será de aplicación y permitirá estimar los resultados globales que se obtendrán en función del método de evaluación que se utilice.

3. Objetivos

Determinar la respuesta estadística de las calificaciones de los estudiantes a lo largo de diferentes cursos, al disponer de una muestra lo suficientemente amplia y valorar si se produce efectivamente en un sistema de evaluación continuada un efecto asimilable al comportamiento de los sistemas susceptibles de ser objeto de la ley de los grandes números y del teorema del límite central.

Comparar dichos efectos tomando sólo las calificaciones debidas a tres exámenes como únicos elementos para obtener el número de aprobados de la asignatura y contrastarlos con todos los elementos de evaluación de la asignatura, esto es, todas las calificaciones obtenidas por cada estudiante utilizando el polinomio de calificación diversificado antes descrito y la técnica de recuperación de suspensos.

4. Desarrollo

De los 3909 estudiantes que han pasado por la asignatura, se ha tomado una muestra aleatoria de 1872 pertenecientes a diferentes cursos académicos desde 2010 hasta 2017, y se ha analizado el valor de sus calificaciones finales, en la que de cada estudiante se dispone de 15 notas correspondientes a ejercicios, 8 correspondientes a prácticas, 3 correspondientes a exámenes escritos, 1 a proyectos y 4 a actividades relacionadas con competencias genéricas. Ello arroja un total de unas 31 calificaciones por estudiante, lo que permite obtener la calificación final. Esta cantidad de alrededor de 30 valores (en función del calendario académico se pueden haber realizado más o menos ejercicios), son calificaciones con valor aleatorio y, por tanto, pueden tomarse como variables aleatorias, aunque no independientes puesto que

poseen una relación entre ellas asociada a la dificultad de los ejercicios realizados y a la posibilidad de recuperación de suspensos. El concepto de valor aleatorio debe entenderse desde el punto de vista de variable estadística, no como que la calificación se establezca de forma aleatoria, sino que el conjunto de calificaciones son un conjunto de datos completamente descorrelacionados entre ellos.

También se han aislado las calificaciones de los 3 exámenes escritos como si todas las calificaciones que se tuvieran de cada estudiante fueran estas y no se tratase de un sistema de evaluación continuada sino de un sistema esencialmente basado en exámenes.

5. Resultados

Se puede ver en la Figura 3 el resultado gráfico de haber tomado solamente los datos correspondientes a 3 exámenes escritos.

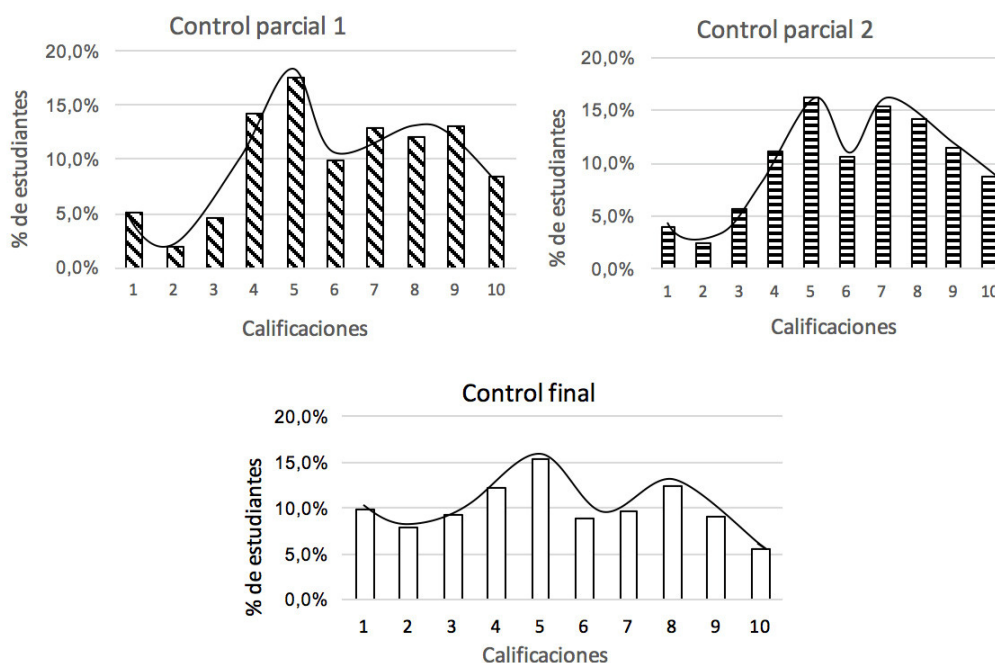


Figura 3: Calificaciones obtenidas en base a considerar sólo los tres exámenes

En dicha figura se puede observar que no se obtiene una forma gaussiana sino más bien unas curvas binodales con picos en la calificación de 5 y 8. El número de aptos sería del 57,9% de los estudiantes evaluados. Esta situación claramente segmenta a los estudiantes aprobados en dos zonas, la de los justamente aptos y la de los notables. Muy pocos son excelentes, sólo un 5% del total. Debe señalarse que entre el 10% de los estudiantes que presentan la calificación más baja, 7 de cada 10 son no presentados, con lo que, realmente, la curva que se obtiene es de naturaleza netamente binodal y el repunte en la zona inferior no debe considerarse.

Este análisis que sólo considera las calificaciones obtenidas en base a tres exámenes, y se calcula cómo se distribuyen las calificaciones con un cálculo de distribución normal, se obtienen los resultados que se pueden ver en la Figura 4. Nótese que el cálculo de una distribución normal, devuelve una distribución normal que, aunque parezca una redundancia, no detalla la incidencia de cada una de las componentes de la calificación (como sí se hace en la Figura 3) sino que las asimila a una forma gaussiana.

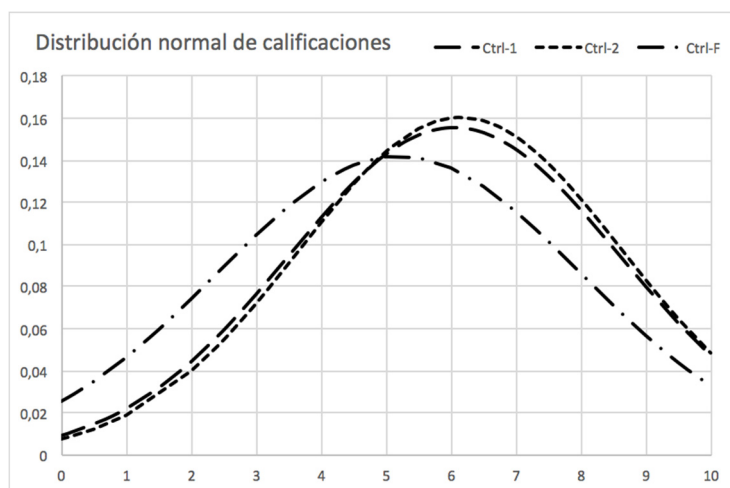


Figura 4: Distribución de resultados debidos a exámenes

En contraste, se puede ver en la Figura 5 el resultado gráfico de haber tomado las alrededor de 30 calificaciones obtenidas por los estudiantes dentro de un sistema de evaluación continuada como el propuesto.

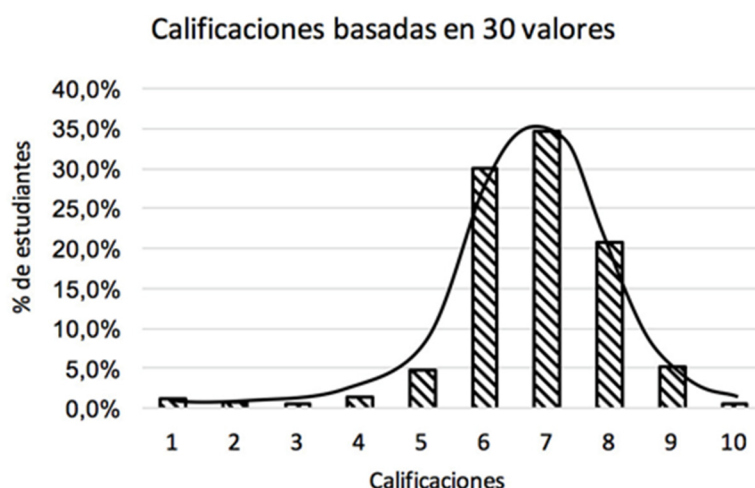


Figura 5: Calificaciones obtenidas en base a 30 calificaciones por estudiante utilizando el polinomio de calificación diversificado

En la Figura 5 se puede observar que se obtiene una forma mucho más gaussiana en vez de binodal como en el caso anterior, con un único máximo situado sobre la calificación de, aproximadamente, notable 7. El número de aptos, en esta ocasión, es del 91,2%, considerando que alrededor de 2% son no presentados, estudiantes que abandonan la asignatura a las pocas semanas de haber empezado. El número de excelentes es de alrededor del 0,6%.

Si sólo se analiza la parte de calificación de las actividades basadas en ejercicios hechos en equipo en clase e individualmente en casa, tomadas en intervalos de 0,5 puntos sobre un colectivo de 630 estudiantes (pertenecientes a cursos en los que se han realizado más ejercicios y por tanto, los datos serán más extensos), se obtiene la distribución de la Figura 6 que tampoco es binodal y tiene un máximo único en la zona entre las calificaciones de 5 y de 7, con media en 5,7.

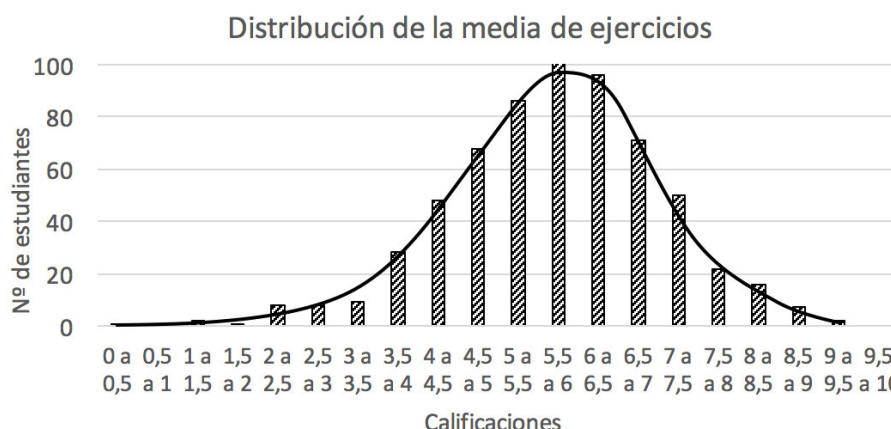


Figura 6: Calificaciones obtenidas por 630 estudiantes en base a 15 calificaciones cada uno, debidas sólo a la parte de ejercicios

Si se hace el análisis de las notas debidas sólo a las calificaciones obtenidas en base a la nota media de los ejercicios, y se calcula cómo se distribuyen estas calificaciones con un cálculo de distribución normal, se obtienen los resultados que se pueden ver en la Figura 7 donde sólo se puede deducir que la media está en 5,7. Como en el caso de la Figura 4, el cálculo de la distribución normal no detalla la incidencia de cada uno de los valores de la calificación en la curva sino que las distribuye en una gráfica normal.

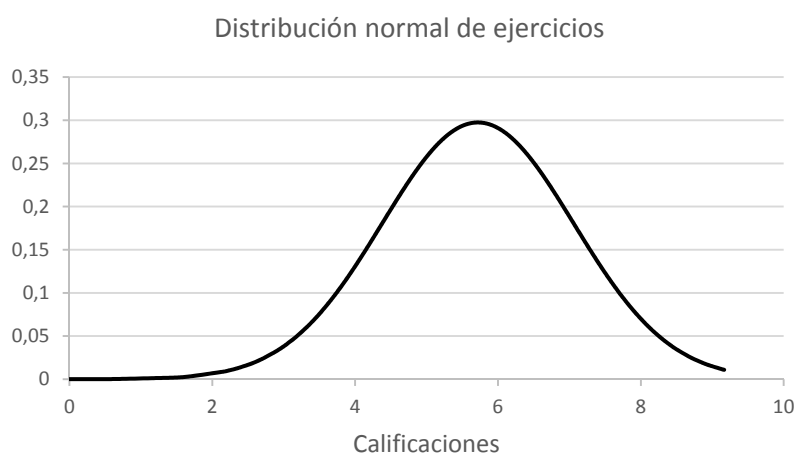


Figura 7: Distribución de resultados debidos a ejercicios

En la Figura 8 puede verse de forma gráfica la diferencia entre ambos modelos de evaluación, el de pocas evaluaciones/calificaciones y el de evaluaciones/calificaciones múltiples. Se observa que en el caso del modelo basado en pocas evaluaciones/calificaciones se asimila a una respuesta binodal con picos en las calificaciones de 5 y 8 aproximadamente, mientras que en el caso de la evaluación con múltiples actividades con calificaciones, se asimila a una respuesta normal desplazada sobre la calificación promedio de aproximadamente 7.

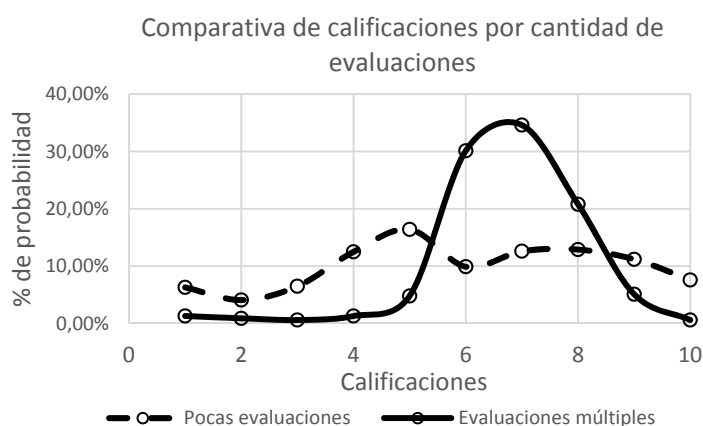


Figura 8: Comparación entre la probabilidad de obtener una determinada calificación en función del número de actos de evaluación realizados

Se han analizado asimismo (véase Figuras 9 y 10) las calificaciones promedio de la parte de 6 prácticas obtenidas por 1432 estudiantes observándose una calificación global media de 6,6 en un cálculo de distribución en la Figura 9 y de distribución normal en la Figura 10 de las mismas.

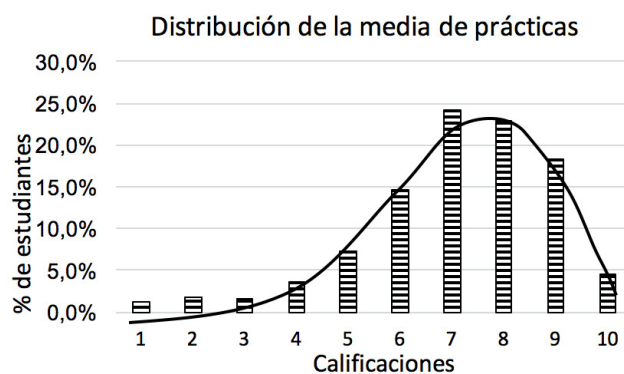


Figura 9: Distribución de resultados debidos a prácticas

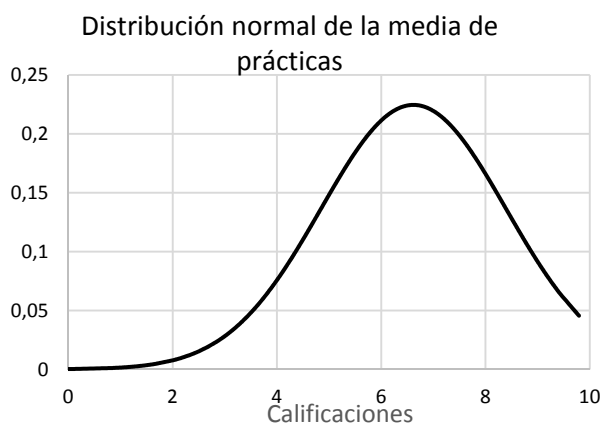


Figura 10: Distribución normal de resultados debidos a prácticas

Puede observarse una cierta asimilación a una distribución gaussiana desplazada a la derecha en la Figura 9.

5.1. Mecanismo de corrección propuesto

Tomando la ley de los grandes números podríamos establecer una analogía por la que la probabilidad de que un estudiante obtenga una determinada calificación, es la de la Tabla 1, donde se compara el sistema basado en pocos exámenes, en evaluación diversificada múltiple y en esta última con un sistema de corrección añadido que se detalla a continuación.

Calificación	Probabilidad en % en un sistema basado en 3 exámenes	Probabilidad en % en un sistema basado en evaluación múltiple	Probabilidad en % en un sistema basado en evaluación múltiple corregida
0-1	6,3%	1,3%	1,3%
1-2	4,1%	0,9%	0,9%
2-3	6,5%	0,6%	0,6%
3-4	12,5%	1,3%	1,3%
4-5	16,4%	4,8%	4,8%
5-6	9,9%	30,1%	30,1%
6-7	12,6%	34,6%	28,5%
7-8	12,9%	20,8%	12,3%
8-9	11,2%	5,1%	13,5%
9-10	7,6%	0,6%	6,8%

Tabla 1: Probabilidad de obtener una determinada calificación según el número de calificaciones debidas a actos de evaluación diferentes

Se observa en la Tabla 1 que mientras que un sistema basado en pocas calificaciones permite que un $11,2\% + 7,6\% = 18,8\%$ de estudiantes alcance las calificaciones superiores mientras que en un sistema basado en calificaciones múltiples, sólo un $5,1\% + 0,6\% = 5,7\%$ las pueden alcanzar.

En la misma Tabla 1 puede verse en la última columna que si se incorpora una corrección, se asimilará a la parte alta de los resultados que se habrían obtenido con un sistema basado en pocas calificaciones, corrigiendo así el efecto ocasionado por no poder obtener siempre calificaciones elevadas y que se ha podido comprobar en la Figura 5; se obtiene con esta corrección que un $13,5\% + 6,8\%$ también alcance las calificaciones superiores.

Para ello se establece una función de tipo polinomial que permite corregir el efecto derivado de tener múltiples pruebas de evaluación con calificaciones, que revierte en la dificultad de poder alcanzar siempre calificaciones elevadas, algo que los datos de que disponemos ponen en evidencia que sí sería posible para los estudiantes más brillantes si se redujera drásticamente el número de pruebas. El sistema de múltiples elementos de evaluación con posibilidad de recuperación favorece a los estudiantes menos brillantes y más lentos en aprendizaje, pero perjudica a los más brillantes, de forma que se toma una función de corrección como la que se presenta en la Figura 11 que permite reescalar las calificaciones de los estudiantes.

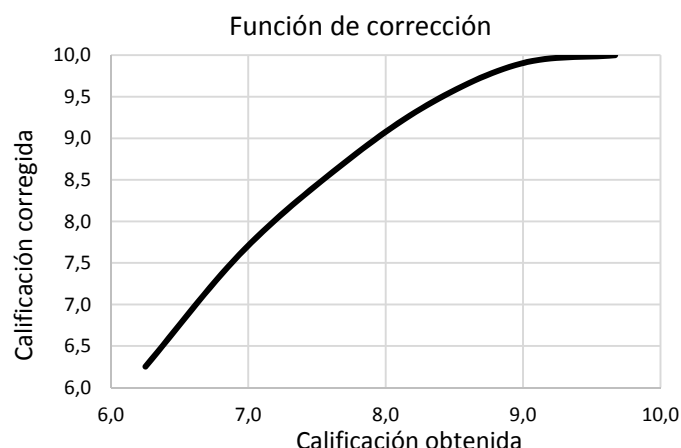


Figura 11: Función de corrección

Esta función se aplica a aquellos estudiantes cuya calificación esté medio punto por encima de la media, así, si la media de las calificaciones finales de todos los estudiantes del curso fuese de, por ejemplo, 6,3, se aplicaría a todos los estudiantes que obtuviesen una media final global de $6,3 + 0,5 = 6,8$. Podría usarse otro criterio puesto que este es arbitrario. Con ello se pretende reequilibrar el resultado que habrían obtenido los estudiantes que estuviesen por encima de dicho valor, en caso de que el número de calificaciones fuera inferior.

Esta corrección se puede ver, a título comparativo en la Figura 12 donde se puede ver el efecto de corrección que la función provoca en la parte alta de las calificaciones donde los estudiantes se desplazan hacia valores superiores de calificación final.

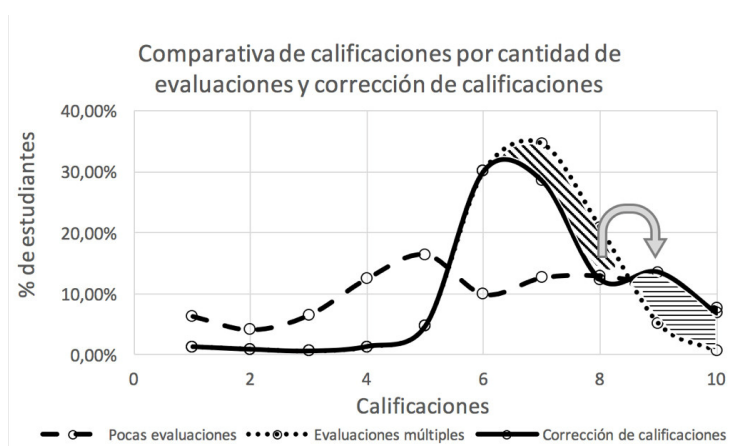


Figura 12: Efecto de la función de corrección del resultado que se obtendría con el método de evaluaciones múltiples

La función utilizada en este caso ha sido $N_f = -0,33N_i^2 + 6,4N_i - 21$ donde N_i es la nota inicial y N_f la nota final corregida. La función en este caso es una cuadrática, pero podría ser lineal o polinomial también, de otro orden superior a 2, a criterio de cada docente. La obtención de esta ecuación es sencilla mediante Excel usando las funciones MINVERSA para calcular la matriz inversa del sistema de ecuaciones de tres variables, en este caso, y MMULT para determinar el valor de dichas tres variables. La obtención del sistema se basa en el cálculo de una regresión polinomial, en este caso, de orden 2. Otra posibilidad reside en añadir, sencillamente, una curva de tendencia a la obtenida gráficamente en Excel, que sea la

representación de la curva de puntos deseada y obtener así una buena aproximación de la función correctora.

Los puntos sobre los que efectuar la regresión son la media de todas las calificaciones obtenidas por todos los estudiantes, que se usa como cota inferior, y como cota superior la máxima obtenida por el estudiante con mayor calificación. Entre ambos valores se establece la forma de la curva de corrección que se desee y con la ayuda de Excel, por ejemplo, se obtiene la función de corrección.

5.2. Discusión

A partir de los resultados obtenidos, se puede llegar a algunas conclusiones en cuanto al comportamiento de un sistema basado en pocas calificaciones y en uno basado en múltiples evaluaciones/calificaciones, así como la corrección del caso de calificaciones múltiples:

- a) El número de no presentados disminuye drásticamente en el sistema de evaluaciones/calificaciones múltiples (estudiantes con calificación en el margen 0-1).
- b) La probabilidad de suspender disminuye drásticamente en el sistema de evaluaciones/calificaciones múltiples ya sea con o sin corrección de la parte alta de calificaciones.
- c) Se comprueba que el teorema del límite central, por el que la distribución de las notas medias en el caso de una suficiente cantidad de estudiantes calificados, tanto si se tienen pocas calificaciones como muchas calificaciones, siempre acaban distribuyéndose como una distribución gaussiana y ello no informa bien a fin de obtener conclusiones relacionadas con la comparación entre sistemas.
- d) Si se ha comprobado que el teorema del límite central en el caso de un solo tipo de actividad, como puedan ser ejercicios o prácticas, cuando se tienen múltiples calificaciones, la distribución de las notas promedio se aproximan razonablemente bien a una distribución normal o, a lo sumo, a una normal desplazada.
- e) Que el número de calificaciones en la franja más alta es superior en el caso de tener pocas calificaciones que en el caso de tener múltiples calificaciones pero que esta situación puede resolverse con una función de corrección que desplace a los estudiantes con calificaciones superiores a la media hacia la parte alta.
- f) Que se puede establecer a priori y siempre que el método de enseñanza/aprendizaje (pocas o múltiples calificaciones) no varíe significativamente, cuál va a ser la probabilidad de obtener una determinada calificación final en función de dicho método.

6. Conclusiones

Se ha establecido que el teorema del límite central no es aplicable a la predicción de los resultados académicos por cuanto siempre asimilará a una distribución normal una muestra suficientemente elevada de calificaciones pero que, no obstante, si se tienen suficientes calificaciones de un estudiante, la distribución de las mismas, se asemeja razonablemente a una distribución normal aunque no informa de la probabilidad real de obtener una determinada calificación; asimismo, se establece que la ley de los grandes números sí aporta información relevante sobre la probabilidad de obtener una determinada calificación y que el uso de una función correctora permite que los estudiantes más brillantes puedan alcanzar las calificaciones más elevadas, corrigiendo el efecto que se da cuando existen una gran cantidad de actos de evaluación que conlleven calificaciones.

Es evidente que se pueden hacer una gran cantidad de evaluaciones formativas, que no tengan una componente sumativa y, en consecuencia no conlleven calificaciones. No obstante, estimamos importante que la valoración de la evaluación formativa se complemente con una cifra que exprese, de alguna manera, en qué cantidad se han alcanzado unos determinados objetivos formativos. Además, disponer de un

polinomio de evaluación diversificado, en el que cada una de sus componentes no tenga un peso determinante, favorece el rendimiento académico de los estudiantes.

La evaluación continuada, con calificaciones asociadas a la misma, mantiene permanentemente informado al estudiante sobre su progreso, y si se le ofrece la posibilidad de aprender de sus errores y mejorar sus productos académicos, entonces, su predisposición al aprendizaje (motivación) aumenta, reduciéndose drásticamente el nivel de abandonos.

Referencias¹

Alonso, C.; Gallego, D. y P. Honey (1994) *Los Estilos de Aprendizaje: Procedimientos de diagnóstico y mejora*. Bilbao: Ediciones Mensajero.

Barnett, R. (2001) *Los límites de la competencia. El conocimiento, la educación superior y la sociedad*. Gedisa.

Boon, J. y M. Van der Klink (2002) The investigation of competencies within professional domains. *Revista Human Resource Development International*, 5(4), pp. 411-424.

del Rey, A. y J. Sánchez-Parga (2011) Crítica de la educación por competencias. *Revista Universitas, Revista de Ciencias Sociales y Humanas*, 15, pp. 233-246.

Delgado, A.M. y R. Oliver (2006) La evaluación continua en un nuevo escenario docente. *Revista de Universidad y Sociedad del Conocimiento*, 3(1), pp. 1-13.

DeSeCo, (2000 y 2005) Es un programa de la OCDE iniciado en 1997 que significa *Definition and Delection of Competencies*. Consultadas en diciembre de 2017 en <http://www.oecd.org/pisa/35070367.pdf> y en <https://www.oecd.org/edu/skills-beyond-school/41529556.pdf>

Díaz Barriga, A. (2006) El enfoque de competencias en la educación. ¿Una alternativa o un disfraz de cambio? *Perfiles Educativos*, XXVIII(111), pp. 7-36.

García Cué, J.L. (2006) *Los Estilos de Aprendizaje y las Tecnologías de la Información y la Comunicación en la Formación del Profesorado*. Tesis Doctoral. Dirigida por Catalina Alonso García. Madrid: Universidad Nacional de Educación a Distancia.

Gardner, H. (1983) *Frames of mind: the theory of multiple intelligences*. New York: Basic Books.

Hervás Avilés, R.M. (2008) Identificación de variables que influyen en los estilos de aprendizaje. Claves para conocer cómo aprenden los estudiantes. *Revista de Estilos de Aprendizaje*, 1(1), pp. 143-167.

Jorba, J. y N. Sanmartí (1994) *Enseñar, aprender y evaluar: un proceso de regulación continua*. Documentos. Ministerio de Educación y Cultura. Barcelona 1994.

Moreno, P. y G. Soto (2005) *Una mirada reflexiva y crítica al enfoque por competencias*. Número de octubre-diciembre de 2005 de la revista Educar.

Visser, B.A.; Ashton, M.C. y P.A. Vernon (2006a) Beyond g: Putting multiple intelligences theory to the test. *Intelligence*, 34(5), pp. 487-502.

Visser, B.A.; Ashton, M.C. y P.A. Vernon (2006b) g and the measurement of Multiple Intelligences: A response to Gardner. *Intelligence*, 34(5), pp. 507-510.

Willis, M. y V.K. Hodson (Eds.) (1999) *Discover your child's Learning Styles*, CA: Prima Publishing.

¹ Alguna de la bibliografia reseñada es deliberadamente antigua puesto que el tema que tratamos no es ninguna novedad en cuanto a sus fundamentos.